Draft version October 5, 2009
Preprint typeset using LaTeX style emulateapj v. 08/22/09

THE EDDINGTON LIMIT IN COSMIC RAYS: AN EXPLANATION FOR THE OBSERVED LACK OF LOW-MASS RADIO-LOUD QUASARS AND THE $M_\bullet - M_\star$ RELATION

Lorenzo Sironi and Aristotle Socrates*
Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544-1001
Draft version October 5, 2009

ABSTRACT

We present a feedback mechanism for supermassive black holes and their host bulges that operates during epochs of radio-loud quasar activity. In the radio cores of relativistic quasar jets, internal shocks convert a fraction of ordered bulk kinetic energy into randomized relativistic ions, or in other words cosmic rays. By employing a phenomenologically-motivated jet model, we show that enough 1-100 GeV cosmic rays escape the radio core into the host galaxy to break the Eddington limit in cosmic rays. As a result, hydrostatic balance is lost and a cosmic ray momentum-driven wind develops, expelling gas from the host galaxy and thus self-limiting the black hole and bulge growth. Although the interstellar cosmic ray power is much smaller than the quasar photon luminosity, cosmic rays provide a stronger feedback than UV photons, since they exchange momentum with the galactic gas much more efficiently. The amount of energy released into the host galaxy as cosmic rays, per unit of black hole rest mass energy, is independent of black hole mass. It follows that radio-loud jets should be more prevalent in relatively massive systems since they sit in galaxies with relatively deep gravitational potentials. Therefore, jet-powered cosmic ray feedback not only self-regulates the black hole and bulge growth, but also provides an explanation for the lack of radio-loud activity in relatively small galaxies. By employing basic known facts regarding the physical conditions in radio cores, we approximately reproduce both the slope and the normalization of the $M_\bullet - M_\star$ relation.

Subject headings: cosmic rays — galaxies: formation — galaxies: fundamental parameters — galaxies: jets

1. INTRODUCTION

There are two classes of persistent sources at cosmic distances: galaxies and quasars/Active Galactic Nuclei (AGN). Stars power galaxies while accretion onto and/or spin down of supermassive black holes power quasars. Until recently, galactic and quasar phenomena were thought to be separate on both observational and theoretical grounds. However, the discovery of the black hole mass – bulge stellar mass relation ($M_\bullet - M_\star$) in nearby elliptical galaxies (Magorrian et al. 1998; McLure & Dunlop 2002; Marconi & Hunt 2003; Häring & Rix 2004)

$$M_\bullet \simeq 10^{-3} M_\star \tag{1}$$

and the black hole mass – stellar velocity dispersion relation ($M_\bullet - \sigma_\star$; Gebhardt et al. 2000; Ferrarese & Merritt 2000; Tremaine et al. 2002)

$$M_\bullet \propto \sigma_\star^4 \tag{2}$$

indicates that galactic and black hole activity are closely connected to one another. The natural implication is that the energy release resulting from the build-up of the black hole mass limits any further growth of both the stellar bulge and the black hole. The fact that the $M_\bullet - M_\star$ relation holds for nearly four decades in black hole mass seems to suggest that a *universal*, *self-similar* or *scale-free* process is at work, which acts to self-regulate the ratio between black hole and bulge mass, irrespective of their combined mass. Apparently, the only question that remains is with regard to the exact physical mechanism responsible for black hole feedback and self-regulation.

The Soltan (1982) argument, along with the work of Yu & Tremaine (2002), indicates that the mass of supermassive black holes is mostly accrued during an optically-luminous radiatively-efficient "quasar phase." The energy released during the accretion process, which is carried away primarily by photons and/or a "quasar wind," may couple to the interstellar medium of the host galaxy and eject it from the galactic gravitational potential (e.g., Silk & Rees 1998; Fabian 1999; Ciotti & Ostriker 2001; King 2003; Di Matteo et al. 2005; Murray et al. 2005; Hopkins et al. 2006). In doing so, fuel for any further galactic and quasar activity is removed and the mass of the black hole, as well as of the stellar bulge, is self-limited.

The energy released during the accretion process may be carried away not solely in the form of photons. In the so called "radio-loud" (as opposed to "radio-quiet") objects, relativistic collimated outflows, or "jets," put out a significant amount of energy in mechanical form. Although radio-loud phenomena are also observed in objects that are not actively accreting (e.g., Ho & Peng 2001; Ho 2002), the radio jet is more likely to affect the evolution of the system when a significant amount of mass is being built up. As this work focuses on the self-regulation of black hole growth, which occurs at high accretion rates, our attention rests on objects that are both significantly accreting and radio-loud.

The kinetic power of radio jets is dissipated in sub-pc scale "radio cores" and kpc to Mpc scale "radio lobes," with comparable amounts of energy dissipated at each

*Current Address: Institute for Advanced Study, Princeton, NJ 08540
Electronic address: lsironi@astro.princeton.edu; socrates@ias.edu

site. The radio-loud quasar phase could be responsible for black hole self-regulation if the energy release from the radio core, unlike that from the distant radio lobes, has the opportunity to couple to the interstellar medium of the host galaxy.

In this work, we propose that black hole self-regulation may be mediated by 1-100 GeV protons, or in other words *cosmic rays*, produced in the jet core during phases of radio-loud quasar activity. We show that enough cosmic rays escape the radio core into the host galaxy to power a cosmic ray-driven wind that ejects the interstellar gas, thus removing the fuel for further star formation and black hole accretion.

Our arguments are organized as follows. In §2 we lay out the basic idea behind our jet-powered cosmic ray feedback scenario by defining and recognizing the importance of the Eddington limit in cosmic rays. In §3 we briefly summarize the physical features of the standard model of radio-loud AGN jets that are important for our feedback mechanism. With the jet model in hand, in §4 we determine the interstellar cosmic ray luminosity resulting from radio-loud quasar activity in terms of observable quantities. For readers uninterested in the details of our jet and radio core model, as well as its phenomenological underpinnings, we suggest skipping §3 and §4. In §5 we examine the consequences of our cosmic ray feedback scenario and make a critical comparison of the black hole self-regulation model presented here with other models of quasar feedback. We summarize our findings in §6, where we also present a brief review of well-known observations that lend support to our jet-powered self-regulation mechanism. In addition, we propose an observational test.

2. THE EDDINGTON LIMIT IN COSMIC RAYS

If supermassive black hole growth results primarily from radiatively-efficient accretion, as the Soltan (1982) argument suggests, then the ratio of the total energy released during the accretion process to the binding energy of the galaxy's gaseous phase is given by

$$\frac{\Delta E_\bullet}{E_g} \simeq \frac{\epsilon_{\rm rad}\, M_\bullet c^2}{f_g M_\star \sigma_\star^2} \simeq 10^3 \left(\frac{M_\bullet/M_\star}{10^{-3}}\right) \epsilon_{\rm rad,\text{-}1}\, f_{g,\text{-}1}^{-1}\, \sigma_{\star,300}^{-2} \ , (3)$$

in case the black hole – galaxy system follows the $M_\bullet - M_\star$ relation in eq. (1). Here, $\epsilon_{\rm rad} = 0.1\,\epsilon_{\rm rad,\text{-}1}$ is the radiative efficiency of the accretion flow. The binding energy of the gas $E_g \simeq f_g M_\star \sigma_\star^2$ is appropriate for a galaxy resembling an isothermal sphere whose gravitational force is mainly provided by stars and dark matter; $f_g = 0.1\, f_{g,\text{-}1}$ is the galaxy's gas fraction and $\sigma_\star = 300\,\sigma_{\star,300}$ km s^{-1} is the stellar velocity dispersion. Clearly, only a small amount of the photon energy released during optically-bright quasar epochs is needed to couple to the galactic gas in order to eject it from the galaxy's gravitational potential, thus self-limiting the combined growth of the black hole and the stellar bulge.

Again, during radio-loud phases a significant amount of the energy resulting from black hole accretion is released in mechanical form as powerful relativistic jets. Even if the time-integrated kinetic output of the jet is $\Delta E_{\rm J} \ll \Delta E_\bullet$, eq. (3) indicates that there may be ample energy to unbind the galactic gas, provided that the fraction of jet kinetic energy deposited in the radio core can efficiently couple to the gaseous component of the host galaxy. In other words, radio-loud kinetically-dominated epochs may represent another mode of AGN activity, other than optically-luminous radiatively-efficient quasar phases, that could potentially lead to efficient black hole self-regulation.

The photon luminosity of the sub-pc scale jet core results from the cooling of relativistic electrons through a combination of synchrotron and inverse Compton emission. By modeling the spectral energy distribution of powerful radio-loud quasars, Celotti & Ghisellini (2008) infer that the emitting electrons must be accelerated, presumably by some magnetized Fermi mechanism, up to highly-relativistic energies, in excess of a few tens of GeV (in the jet comoving frame). The details of the acceleration mechanism aside, it is reasonable to conclude as well that protons are randomized with energies up to $\sim 1-10$ GeV. If a significant fraction of these randomized relativistic ions, or in other words *cosmic rays*, escape the site of acceleration and diffuse into the interstellar medium of the host galaxy, they may provide the coupling between jet power and galactic gas required for black hole self-regulation during radio-loud phases.

For example, in actively star-forming galaxies the generation and subsequent diffusion of cosmic ray protons – a well-known by-product of core-collapse supernovae, and thus massive star-formation – may act as an agent of self-regulation, limiting the rate of star-formation and therefore the galaxy luminosity. The simplest way to understand this is by defining the *Eddington limit in cosmic rays*

$$L_{\rm Edd,CR} = L_{{\rm Edd},\star}\frac{\lambda_{\rm CR}}{\lambda_{\rm T}} \simeq 1.3\times10^{44}\left(\frac{\lambda_{\rm CR}/\lambda_{\rm T}}{10^{-6}}\right)\left(\frac{M_\star}{10^{12}M_\odot}\right)\,{\rm erg\,s^{-1}} \ , (4)$$

where $L_{{\rm Edd},\star}$ is the Thomson Eddington limit for a galaxy with mass $M_\star$; we have pinned the ratio between the cosmic ray mean free path and the Thomson mean free path to the value $\lambda_{\rm CR}/\lambda_{\rm T} \sim 10^{-6}$, as in the case of the Milky Way. An interstellar cosmic ray luminosity of $L_{\rm CR} \sim 10^{44}\,{\rm erg\,s^{-1}}$ – equal to the corresponding cosmic ray Eddington limit $L_{\rm Edd,CR}$ for the most massive galaxies – may result from the act of forming stars at a rate of $\sim 10^3\, M_\odot\,{\rm yr^{-1}}$ – corresponding to the brightest star-forming galaxies. The formation of stars at a higher rate leads to the breaking of the cosmic ray Eddington limit and a cosmic ray-driven wind develops, removing the gaseous phase of the galaxy thus choking off star-formation. From this straightforward argument, it is quite possible that the luminosities of star-forming galaxies are capped at their observed values because they are Eddington-limited in cosmic rays, as proposed by Socrates et al. (2008).

With respect to AGN, radio-loud objects display powerful jets with kinetic luminosity upwards of $L_{\rm J} \sim 10^{47}\,{\rm erg\,s^{-1}}$ (e.g., Rawlings & Saunders 1991), comparable to the Thomson Eddington limit $L_{{\rm Edd},\bullet}$ for the most massive black holes with $M_\bullet \sim 10^9\,M_\odot$. If the jet kinetic power is a fraction $\Lambda_{\rm Edd} \sim 1$ of the black hole Thomson Eddington limit, we have

$$L_{\rm J} = \Lambda_{\rm Edd} L_{{\rm Edd},\bullet} \simeq 1.3\times10^{47}\Lambda_{\rm Edd} M_{\bullet,9}\,{\rm erg\,s^{-1}} \ , \quad (5)$$

where $M_\bullet = 10^9 M_{\bullet,9}\, M_\odot$. Let $\epsilon_{\rm CR}$ be the efficiency for converting the jet kinetic power $L_{\rm J}$ into interstellar cosmic ray power $L_{\rm CR}$. A statement of momentum conservation and hydrostatic balance, the Eddington limit in cosmic rays defined in eq. (4) indicates that a galactic cosmic ray-driven wind develops if

$$L_{\rm CR} = \epsilon_{\rm CR}\, L_{\rm J} \gtrsim L_{\rm Edd,CR}\ , \tag{6}$$

or if the fraction of jet kinetic luminosity ending up as interstellar cosmic rays attains a value of

$$\epsilon_{\rm CR} \gtrsim 10^{-3}\, \Lambda_{\rm Edd}^{-1} \left(\frac{\lambda_{\rm CR}/\lambda_{\rm T}}{10^{-6}}\right) \left(\frac{M_\star/M_\bullet}{10^3}\right) \tag{7}$$

for a black hole – galaxy system that follows the $M_\bullet - M_\star$ relation.

The *momentum* requirement in eqs. (6) or (7) is a necessary condition to initiate a galactic cosmic ray-driven outflow. However, in order to fully unbind the galaxy's gaseous phase, thus self-regulating the black hole and bulge growth, this epoch of super-Eddington cosmic ray activity should last long enough so that the time-integrated *energy* injected as cosmic rays into the interstellar medium is comparable to the binding energy of the galactic gas. Interestingly, the value of the jet-cosmic ray conversion efficiency given in eq. (7) is roughly equal to $E_g/\Delta E_\bullet$ for the most massive galaxies (see eq. (3)), where $E_g/\Delta E_\bullet$ may be viewed as the minimum efficiency of a "generic" feedback mechanism for coupling the black hole energy release to the gaseous component of the host galaxy. In other words, the minimum value of $\epsilon_{\rm CR}$ required to break the galaxy's *momentum* balance, given in eq. (7), is approximately equal to the minimum value of the *energy* coupling efficiency $E_g/\Delta E_\bullet$ for the most massive systems.[1]

Clearly, an understanding and determination of the jet-cosmic ray efficiency parameter $\epsilon_{\rm CR}$ is of fundamental importance with respect to quantifying whether or not jet-cosmic ray feedback is in fact the agent of self-regulation for a given black hole – galaxy system. In what follows, we take into account the physical properties of AGN radio cores with the help of a well-established jet model. In doing so, we are able to determine the jet-cosmic ray efficiency parameter $\epsilon_{\rm CR}$ and discuss the effectiveness of our cosmic ray-driven feedback scenario.

3. A STANDARD JET MODEL FOR RADIO-LOUD AGN

Here we describe the physical properties of quasar jets, whose radio core is responsible for injecting cosmic rays into the interstellar medium of the host galaxy. The reasoning and details of the jet model outlined below are grounded by decades of observations of radio-loud quasars. We also discuss some of the observational constraints that we use to determine the parameters of our model.

3.1. *Basics of Radio-Loud AGN Phenomena*

The spectral signature of radio-loud AGN results from dissipation of the kinetic energy of powerful relativistic jets in sub-pc scale "radio cores" and kpc to Mpc scale "radio lobes." The basic radio-loud phenomenology is largely understood in terms of a viewing-angle effect (e.g., Antonucci 1993; Urry & Padovani 1995). Emission from the relativistic jet flow in the radio core is highly beamed due to relativistic aberration. The radiation seen along the jet axis is Doppler-boosted in both frequency and flux, when compared to the flow rest frame, and the opposite is true when the radio core is viewed edge-on.

Core-dominated on-axis sources are collectively referred to as "blazars" – a marriage between BL Lacs and flat-spectrum radio quasars (FSRQs). They display extreme variability, broad (radio to gamma-ray) spectral energy distributions, super-luminal motion and super-Eddington fluxes. BL Lac spectra are typically featureless, whereas intense line emission is detectable in the optical-UV spectrum of FSRQs, presumably arising from the accretion disk and its associated broad-line region (BLR). The continuum spectral energy distribution of blazars shows two prominent bumps, with the high-energy bump that often dominates the bolometric power output. The low-energy bump, at infrared to UV frequencies, is conventionally attributed to synchrotron radiation, while the high-energy emission, extending up to gamma-ray energies, is thought to result from inverse Compton up-scattering of synchrotron photons (in BL Lacs) or external photons from the accretion disk and the BLR (in FSRQs).

In objects where the jet is viewed edge-on, the so called "Fanaroff-Riley" galaxies, the radio emission mostly results from the dissipation of jet power in radio lobes on inter-galactic or super-galactic scales. The emission from sub-pc scales is largely beamed away from the observer for these off-axis sources. In the framework of unification models (e.g., Antonucci 1993; Urry & Padovani 1995), FR I radio galaxies are identified as the parent population of BL Lacs, while FR II galaxies are the edge-on counterpart of FSRQs.

In this work, all of our attention will be directed towards the most powerful objects, i.e., FSRQs and FR II galaxies, sources where radio-loud activity coexists with substantial black hole accretion. From here on, FSRQs and their parent population of FR II radio galaxies will be collectively referred to as "radio-loud quasars".

3.2. *A Simple Jet Model*

We assume that the jet power or "luminosity" $L_{\rm J}$ is primarily carried in mechanical form by matter expelled from the central engine at a rate $\dot{M}_{\rm J}$, such that

$$L_{\rm J} \simeq \Gamma \dot{M}_{\rm J} c^2\ , \tag{8}$$

where Γ is the characteristic Lorentz factor of the flow. Under the assumption that the ultimate source of jet power is accretion onto the black hole at a rate $\dot{M}_\bullet$ (as opposed to extraction of the black hole spin), then $\dot{M}_{\rm J}/\dot{M}_\bullet \sim \epsilon_{\rm rad}/\Gamma$ if the jet kinetic power is comparable to the accretion photon luminosity, $L_{\rm J} \sim L_{\rm acc} \simeq \epsilon_{\rm rad} \dot{M}_\bullet c^2$. It follows that for typical values, namely $\epsilon_{\rm rad} \sim 0.1$ and $\Gamma \sim 10$ (e.g., Jorstad et al. 2005), we have $\dot{M}_{\rm J}/\dot{M}_\bullet \sim 0.01$. In other words, a significant fraction of the accretion power must be transmitted to only $\sim 1\%$ of the mass. The mechanism of this seemingly improbable transfer of energy is assumed to be an unknown from

[1] Strictly speaking, the minimum energy coupling efficiency for our jet-powered cosmic ray feedback would be $\sim E_g/\Delta E_{\rm J}$, where $\Delta E_{\rm J}$ is the time-integrated kinetic output of the radio jet. We are implicitly assuming that $\Delta E_{\rm J} \sim \Delta E_\bullet$ at the high-mass end.

here on as we are primarily interested in its mechanical output.

Powerful radio-loud quasars are variable, often displaying radio blobs, or "knots," moving along the jet. Such behavior is commonly thought to result from the dissipative interaction of shells of matter intermittently ejected from the central engine. If the Lorentz factors of the shells differ such that $\Delta\Gamma \sim \Gamma$ in the reference frame of the host galaxy (here, Γ may be thought of as the mean Lorentz factor of the jet flow), then the characteristic distance $R_{\rm diss}$ from the black hole where dissipation takes place is

$$R_{\rm diss} \simeq \frac{\Gamma^2}{\epsilon_{\rm rad}} R_{\rm G} \simeq 0.05\ \Gamma_1^2\, \epsilon_{\rm rad,\text{-}1}^{-1} M_{\bullet,9}\ {\rm pc}\ , \qquad (9)$$

where $\Gamma = 10\,\Gamma_1$; $R_{\rm G} = GM_\bullet/c^2$ is the black hole gravitational radius and $R_{\rm G}/\epsilon_{\rm rad} c$ is the dynamical timescale of the accretion flow, which determines the typical time delay between subsequent shell ejections by the central engine. For the sake of simplicity, we adopt a one-zone model in which the whole jet spectral energy distribution originates from a single region located at $R_{\rm diss}$.

We assume that the jet is beamed into a cone with half-opening angle $\theta \sim 0.1$ (e.g., Jorstad et al. 2005). For typical values $\Gamma \sim 10$ and $\theta \sim 0.1$, in the jet comoving frame the dissipation region resembles a spherical blob, since its comoving longitudinal length $\sim R_{\rm diss}/\Gamma$ is comparable to its transverse size $\sim R_{\rm diss}\theta$. For a bi-conical mass-dominated jet,

$$L_{\rm J} \simeq \Gamma\, \dot{M}_{\rm J} c^2 \simeq 2\pi (R\,\theta)^2 \Gamma \rho\, c^3\ , \qquad (10)$$

where R is the distance from the black hole and ρ is the jet mass density at radius R measured in the host-galaxy frame. If the jet kinetic power is a fraction $\Lambda_{\rm Edd}$ of the black hole's photon Eddington limit, the density $\rho_{\rm diss}$ at the dissipation scale $R_{\rm diss}$ and the Thomson optical depth $\tau_{\rm diss}$ down to $R_{\rm diss}$ may be written in terms of basic physical parameters as

$$\rho_{\rm diss} \simeq \frac{2\,\epsilon_{\rm rad}^2 \Lambda_{\rm Edd}}{\Gamma^5 \theta^2 \kappa_{es} R_{\rm G}} \qquad (11)$$

and

$$\tau_{\rm diss} \simeq \frac{2\,\epsilon_{\rm rad}\Lambda_{\rm Edd}}{\Gamma^3\,\theta^2}\frac{n_e}{n_p} \simeq 0.02\,\frac{n_e}{n_p}\,\epsilon_{\rm rad,\text{-}1}\Lambda_{\rm Edd}\Gamma_1^{-3}\theta_{\text{-}1}^{-2}\ , \qquad (12)$$

where $\theta = 0.1\,\theta_{\text{-}1}$ and $\kappa_{es} \simeq 0.4\ {\rm cm^2\,g^{-1}}$ is the electron scattering opacity; n_p and n_e are the jet number densities of protons and leptons (including both positrons and electrons) in the host-galaxy frame. In eq. (12) we assume that the jet kinetic power is mostly carried by protons – in §3.3.2 we estimate that pairs outnumber ions by only a factor of ten. For $n_e/n_p \sim 10$, it follows from eq. (12) that the dissipation region is optically thin.

In the dissipation region, internal shocks resulting from shell-shell collisions convert a fraction $\epsilon_{\rm th}$ of the orderly mechanical energy of the jet flow into heat. It follows that the randomized thermal power generated in the dissipation region is

$$L_{\rm th} = \epsilon_{\rm th} L_{\rm J}\ , \qquad (13)$$

which corresponds to a thermal energy density in the comoving frame

$$U'_{\rm th} = \epsilon_{\rm th}\,\rho'_{\rm diss}\, c^2 = \frac{\epsilon_{\rm th}}{\Gamma}\rho_{\rm diss} c^2\ , \qquad (14)$$

where $\rho'_{\rm diss} = \rho_{\rm diss}/\Gamma$ is the jet mass density at $R_{\rm diss}$ measured in the comoving frame. If the two colliding shells have equal rest mass and differ by $\Delta\Gamma$ in Lorentz factor, the fraction of bulk kinetic energy converted into heat is $\epsilon_{\rm th} \simeq 1/8\,(\Delta\Gamma/\Gamma)^2$ to leading order in $\Delta\Gamma/\Gamma$ (e.g., Kobayashi et al. 1997). We take $\epsilon_{\rm th} \sim 0.1$ as a benchmark for shells with $\Delta\Gamma \sim \Gamma$. It follows that the average comoving Lorentz factor of shocked protons is $\gamma_p \simeq 1+\epsilon_{\rm th}$, corresponding to a mean energy $\simeq \Gamma\,\gamma_p\, m_p c^2 \sim 10$ GeV in the frame of the host galaxy. If a sufficient number of these randomized relativistic ions, or cosmic rays, escape the jet into the galaxy's interstellar medium, they may provide the coupling between jet power and galactic gas required for our feedback model.

Since radio cores are inarguably powered by synchrotron emission, our jet model cannot be complete without an estimate of the magnetic field. We parametrize the magnetic energy density in the comoving frame $U'_{\rm B}$ as a fraction $\epsilon_{\rm B} \sim 0.1$ of the thermal energy density $U'_{\rm th}$:

$$U'_{\rm B} = \epsilon_{\rm B} U'_{\rm th}\ . \qquad (15)$$

It is reasonable to suspect that the magnetic field is tangled and inhomogenous within the dissipation region, more so since we assume the jet electromagnetic flux to be sub-dominant with respect to the kinetic flux. In the comoving frame, the mean magnetic field strength is

$$B' \simeq 2.7\,\epsilon_{\rm B,\text{-}1}^{1/2}\,\epsilon_{\rm th,\text{-}1}^{1/2}\,\epsilon_{\rm rad,\text{-}1}\Lambda_{\rm Edd}^{1/2}\,\Gamma_1^{-3}\,\theta_{\text{-}1}^{-1} M_{\bullet,9}^{-1/2}\ {\rm G}\ , \qquad (16)$$

where $\epsilon_{\rm th} = 0.1\,\epsilon_{\rm th,\text{-}1}$ and $\epsilon_{\rm B} = 0.1\,\epsilon_{\rm B,\text{-}1}$.

Now, we possess all of the necessary physical ingredients for our feedback model. However, before we assess whether or not a sufficient number of cosmic ray protons are able to leak out of the radio core, potentially leading to the disruption of the entire host galaxy, we spend some time – for the sake of completeness – on the observational motivations that went into our jet model.

3.3. *Observational Constraints on the Jet Model*

Many of the basic physical features of the somewhat "canonical" jet model outlined above are strongly supported and motivated by decades of observations. A synopsis of the corroborating evidence follows.

3.3.1. *Jet Kinetic Luminosity: $L_{\rm J}$ and $\Lambda_{\rm Edd}$*

Edge-on Fanaroff-Riley jets are better suited than head-on blazars to measure the jet kinetic power. For blazars, the highly Doppler-beamed, variable, broadband radiation output must be properly modeled in order to extract physical parameters at the jet dissipation scale. At any given frequency, the value of the comoving radiation flux $F'_{\nu'}$ itself is difficult to determine. In fact, in the observer frame,

$$F_\nu \simeq \delta^3 F'_{\nu'}\ , \qquad (17)$$

where δ is the Doppler-beaming factor (e.g., Lind & Blandford 1985), which depends upon Γ and the orientation of the observer's line of sight with respect to the

jet axis. Clearly, due to the strong dependence on the Doppler-beaming factor δ, an accurate measurement of the comoving radiation flux $F'_{\nu'}$, and therefore of intrinsic jet properties such as $L_{\rm J}$ and $\Lambda_{\rm Edd} = L_{\rm J}/L_{\rm Edd,\bullet}$, is difficult to perform.

In the case of Fanaroff-Riley sources, the terminal sites of jet dissipation, the so called radio lobes, can be thought of as the thermal reservoirs of the initially ordered and collimated jet mechanical energy. Since the energy injection rate outstrips the radiative cooling rate, radio lobes have no choice other than to expand. Plasma temperature and density within a radio lobe are typically extrapolated from its X-ray continuum and line emission. From the inferred lobe enthalpy $E_{\rm lobe}$, the jet kinetic power can be estimated as

$$L_{\rm J} \simeq \frac{E_{\rm lobe}}{t_{\rm lobe}} \simeq \frac{E_{\rm lobe}\, c_s}{l_{\rm lobe}} \ , \tag{18}$$

where $t_{\rm lobe} \simeq l_{\rm lobe}/c_s$ is the characteristic time required to inflate a lobe with width $l_{\rm lobe}$ at the sound speed c_s (e.g., Bîrzan et al. 2004; Shurkin et al. 2008).

At the powerful end, the lobes of FR II galaxies possess average kinetic inputs of $L_{\rm J} \simeq 10^{47} - 10^{48}\,{\rm erg\,s^{-1}}$, corresponding to the Thomson Eddington limit for black holes with mass $10^9 - 10^{10}\,M_\odot$.[2] The fact that the upper limit for $L_{\rm J}$ is close to the Eddington limit for the largest black holes informs us that, at least during periods of powerful activity, a value of $\Lambda_{\rm Edd} \simeq 1$ may be appropriate for these sources, and that $\Lambda_{\rm Edd} \sim 1$ may serve as a reasonable upper bound. For powerful FSRQs, thought to be the on-axis counterpart of massive FR II galaxies, Celotti & Ghisellini (2008) infer comparable values for $L_{\rm J}$ from modeling the blazar spectral energy distribution.

3.3.2. *Optical Depth and Pair Content: $\tau_{\rm diss}$ and n_e/n_p*

Information on particle composition at the jet dissipation scale is best extracted from the spectra of FSRQs, where both jet and disk emission are present. The general procedure is outlined by Sikora et al. (1997).

The jet is bathed by optical/UV photons from the BLR and IR photons from the presumed obscuring torus. Owing to the large bulk Lorentz factor $\Gamma \sim 10$ of the jet flow, "cold" electrons in the jet have the capacity to Compton up-scatter these relatively soft photons (a process known as "bulk-Comptonization" or first-order kinetic Sunyaev - Zel'dovich effect) up to characteristic energies

$$h\nu_{\rm BC} \simeq \Gamma^2 h\nu_{\rm UV} \simeq \Gamma_1^2 \ {\rm keV} \ , \tag{19}$$

where $h\nu_{\rm UV} \simeq 10$ eV is the typical seed photon energy. The expected bulk-Compton luminosity $L_{\rm BC}$ from cold jet electrons may be written as a volume integral extending from the jet base $\sim R_{\rm G}/\epsilon_{\rm rad}$ to the dissipation scale $R_{\rm diss}$, i.e., below the region where electrons are shock-heated:

$$\begin{aligned} L_{\rm BC} &\simeq \frac{\delta^3}{\Gamma}\int {\rm d}V n_e \left|\frac{{\rm d}E_e}{{\rm d}t}\right| \simeq 2\Gamma^2 \int \ {\rm d}V n_e\, c\,\sigma_{\rm T}\Gamma^2\, U_{\rm BLR} \\ &= \frac{1}{2}\Gamma^4\theta^2\,\xi\, L_{\rm acc}\tau_{\rm cold} \ , \end{aligned} \tag{20}$$

[2] Black-hole mass is usually estimated from optical and UV lines, since their width supposedly measures the depth of the gravitational potential at the BLR.

where ${\rm d}E_e/{\rm d}t$ is the rate at which a cold electron loses its orderly kinetic energy via the bulk-Comptonization process, $\sigma_{\rm T}$ is the Thomson cross section and $U_{\rm BLR} = \xi\, L_{\rm acc}/(4\pi R^2 c)$ is the energy density at radius R resulting from the fraction ξ of accretion luminosity $L_{\rm acc}$ reprocessed and isotropized by the BLR. Eq. (20) leads to an expression for the Thomson optical depth $\tau_{\rm cold}$ from the jet base up to the dissipation region:

$$\tau_{\rm cold} \simeq \frac{2}{\Gamma^4\,\theta^2}\frac{L_{\rm BC}}{\xi\, L_{\rm acc}} \lesssim 0.2\left(\frac{L_{\rm SX}/\xi L_{\rm acc}}{10}\right)\Gamma_1^{-4}\theta_{\text{-}1}^{-2} \ , \tag{21}$$

where we have imposed that the expected bulk-Compton luminosity $L_{\rm BC}$ should not exceed the observed integrated power $L_{\rm SX}$ at soft X-ray energies (where the bulk-Compton emission should peak, see eq. (19)).

Assuming that the flux of cold electrons is conserved along the jet (i.e., pair injection occurs primarily at the jet base), it follows that the Thomson depth $\tau_{\rm diss}$ down to the dissipation scale satisfies $\tau_{\rm diss} \lesssim \tau_{\rm cold}$, since the optical depth is greater at smaller radii. Thus, from eqs. (12) and (21) the electron to proton ratio of the jet flow is constrained to be

$$\frac{n_e}{n_p} \lesssim 10\left(\frac{L_{\rm SX}/\xi L_{\rm acc}}{10}\right)\Lambda_{\rm Edd}^{-1}\epsilon_{\rm rad,\text{-}1}^{-1}\Gamma_1^{-1} \ , \tag{22}$$

which informs us that electrons and positrons, though dominant by number, only advect with them an insignificant fraction of the jet power. That is, the kinetic power of the jet is almost entirely carried by ions, as assumed in §3.2. Also, from eq. (12) with $n_e/n_p \sim 10$ it follows that $\tau_{\rm diss} \simeq 0.2$, and the flow in the dissipation region is optically thin.

3.3.3. *Magnetic Field Strength: $\epsilon_{\rm B}$*

As in the case of estimating the jet composition, FSRQs are the best sources for measuring the relative strength of the magnetic field at the dissipation scale, parametrized by $\epsilon_{\rm B} = U'_{\rm B}/U'_{\rm th}$. In the event that all of the low-energy synchrotron (with integrated luminosity $L_{\rm S}$) and high-energy inverse Compton ($L_{\rm IC}$) emission is powered by electrons accelerated in the dissipation region, then

$$\frac{L_{\rm S}}{L_{\rm IC}} \simeq \frac{U'_{\rm B}}{U'_{\rm soft}} = \epsilon_{\rm B}\frac{U'_{\rm th}}{U'_{\rm soft}} \ , \tag{23}$$

where $U'_{\rm soft}$ is the comoving energy density of seed photons for the inverse Compton process. For FSRQs, $U'_{\rm soft}$ is mostly contributed by external photons from the BLR and obscuring torus (as opposed to synchrotron seed photons, that dominate in BL Lacs), so that $U'_{\rm soft} \sim \Gamma^2\, U_{\rm BLR}$, where $U_{\rm BLR}$ was defined in §3.3.2. With help from eq. (14) for the form of $U'_{\rm th}$ we have

$$\frac{U'_{\rm th}}{U'_{\rm soft}} \simeq \frac{2\,\epsilon_{\rm th}}{\Gamma^4\theta^2\xi} \ , \tag{24}$$

where we made liberal use of the fact that $L_{\rm J} \sim L_{\rm acc}$ for our jet model in FSRQs. It follows that

$$\epsilon_{\rm B} \simeq 0.05\left(\frac{L_{\rm S}/L_{\rm IC}}{0.1}\right)\left(\frac{\xi}{10^{-3}}\right)\epsilon_{\rm th,\text{-}1}^{-1}\Gamma_1^4\,\theta_{\text{-}1}^2 \ , \tag{25}$$

which is consistent with the value $\epsilon_{\rm B} \sim 0.1$ chosen in §3.2. Comparable values result from detailed modeling

of blazar spectral energy distributions (Celotti & Ghisellini 2008). This confirms that the electromagnetic contribution to the jet energy flux at the dissipation scale is negligible compared to the proton kinetic flux.

The magnetic energy fraction $\epsilon_{\rm B}$ computed in eq. (25) is independent from the black hole mass $M_\bullet$. The apparent lack of dependence of $\epsilon_{\rm B}$ on $M_\bullet$ may be misleading. In fact, the fraction ξ of accretion luminosity that is re-processed and isotropized in the BLR will in principle depend upon $M_\bullet$, since the characteristic disk temperature scales as $\propto M_\bullet^{-1/4}$ at fixed Eddington ratio (Shakura & Sunyaev 1973), which leads to a corresponding change in the photo-ionizing luminosity per unit of accretion power. Nevertheless, for the level of accuracy of this work we ignore such complications.

3.4. *Summary of the Jet Model*

Our "canonical" jet model may be considered more as a "consensus" jet model. That is, it is based upon decades of observations and modeling of radio-loud quasars, rather than any deep theoretical principle. We do not address the nature of the ultimate source of mechanical energy at the base of the jet. We simply exploit the fact that the jet is launched with some mixture of protons and electrons at a constant Lorentz factor $\Gamma \sim 10$ and half-opening angle $\theta \sim 0.1$, as often inferred in bright core-dominated sources. Although pairs may outnumber protons, the jet power is dominated by the ions. The most important feature of the jet model in terms of cosmic ray production is that dissipation by internal shocks at a distance $R_{\rm diss} \simeq \Gamma^2 R_{\rm G}/\epsilon_{\rm rad}$ from the black hole leads – almost certainly on theoretical grounds – to the randomization of bulk baryonic energy and – without doubt on observational grounds – of bulk leptonic energy. The shock-heated protons that escape the jet into the interstellar medium of the host galaxy may have profound implications for the galaxy evolution.

The most important aspect of our jet model is that it is extremely simple with relatively few inputs. That is, we adopt typical values for Γ, $\Delta\Gamma$, θ and $\epsilon_{\rm rad}$ and we derive physically-motivated estimates for $\Lambda_{\rm Edd}$, $\epsilon_{\rm th}$ and $\epsilon_{\rm B}$. From these parameters, we extract all the necessary information required to model our jet-powered cosmic ray feedback mechanism, as we describe in the next section.

4. DETERMINATION OF $\epsilon_{\rm CR}$ AND $L_{\rm CR}$ FOR RADIO-LOUD QUASAR ACTIVITY

In our model of quasar self-regulation, we rely upon interstellar cosmic rays generated in the radio core of quasar jets. As discussed in §3.2, in the jet dissipation region a fraction $\epsilon_{\rm th} \sim 0.1$ of the jet's bulk kinetic power is converted into random thermal form by internal shocks. The average comoving energy of shocked protons will be in the GeV range, since their mean Lorentz factor is $\gamma_p \simeq 1+\epsilon_{\rm th}$; this corresponds in the host-galaxy frame to a kinetic energy $\simeq \Gamma\, \gamma_p\, m_p c^2 \sim 10$ GeV. We now address the question of whether a sufficient number of these *moderately-relativistic* cosmic rays can diffuse out of the jet into the host galaxy, thus providing the coupling between jet power and interstellar gas required for our feedback scenario.

We find that most of the randomized protons produced by internal shocks are convected with the jet flow away from the dissipation region without suffering significant losses. Only a small fraction ($\sim t_{\rm adv}/t_{\rm diff,J}$) can escape from the jet before being advected into regions where magnetic irregularities in the flow are too weak to allow for significant particle diffusion. Here, $t_{\rm adv}$ is the advection time through the dissipation region and $t_{\rm diff,J}$ is the diffusion time across the jet, as measured in the jet comoving frame.[3] It follows that the fraction of jet kinetic luminosity $L_{\rm J}$ ending up as interstellar cosmic ray power will be

$$\epsilon_{\rm CR} \simeq \epsilon_{\rm th} \frac{t_{\rm adv}}{t_{\rm diff,J}}\ . \tag{26}$$

It is the purpose of this section to estimate the cosmic ray efficiency $\epsilon_{\rm CR}$ and the corresponding interstellar cosmic ray power $L_{\rm CR} = \epsilon_{\rm CR} L_{\rm J}$. Comparison of $L_{\rm CR}$ with the Eddington limit in cosmic rays defined in eq. (4) will assess if enough cosmic rays are injected into the host galaxy to initiate a momentum-driven wind capable of removing the galactic gas, thus potentially self-regulating the black hole – galaxy co-evolution.

4.1. *Important Timescales for* $1-10$ GeV *Cosmic Rays in the Jet Dissipation Region*

Table 1 summarizes the important comoving timescales for 1-10 GeV cosmic rays in the jet dissipation region; in boldface, the timescales most relevant for our model. As discussed below, we find that: *i*) randomized relativistic protons are produced by internal shocks on a much shorter timescale than advection, which is itself the fastest loss process for 1-10 GeV ions; *ii*) cosmic ray diffusion out of the jet is a slow process compared to advection, implying that only a small fraction of the shock-accelerated protons will contribute to the interstellar cosmic ray population.

4.1.1. *Cosmic Ray Production and Losses*

Although eq. (12) implies that the dissipation region is collisionless, the presence of even modest magnetic fields reasonably ensures that the transition from the fast cold pre-shock flow to the relatively-slow hot post-shock medium occurs within a few Larmor scales (as observed for the Earth's bow shock by, e.g., Tanaka et al. 1983; Sckopke et al. 1990). In the jet comoving frame, this implies that internal shocks heat and randomize the incoming cold protons on the short timescale

$$t_{\rm heat} \simeq \omega_{\rm L}^{-1} \simeq 3.9\times10^{-5}\,\gamma_p \frac{\Gamma_1^3\,\theta_{\text{-}1}}{\Lambda_{\rm Edd}^{1/2}\,\epsilon_{\rm B,\text{-}1}^{1/2}\,\epsilon_{\rm th,\text{-}1}^{1/2}\,\epsilon_{\rm rad,\text{-}1}}\, M_{\bullet,9}^{1/2}\ {\rm s}\ , \tag{27}$$

where $\omega_{\rm L} = eB'/(\gamma_p m_p c)$ is the Larmor frequency of shock-heated protons with characteristic Lorentz factor γ_p; B' is the comoving magnetic field strength in eq. (16). We remark that our cosmic ray-driven feedback model does not require highly-relativistic protons, possibly accelerated at internal shocks by some Fermi mechanism. Rather, the bulk of the self-regulation effect results from "thermal" protons with comoving energies in the GeV range, produced in the dissipation region on the timescale $t_{\rm heat}$ computed above. For the sake of completeness, we also include in Table 1 the characteristic timescale $t_{\rm DSA}$

[3] We point out that all timescales in this section are computed in the jet comoving frame.

of diffusive shock acceleration.[4]

Production of $\sim$ GeV cosmic rays on the timescale $t_{\rm heat}$ estimated above is much faster than advection through the dissipation region, which happens on a characteristic time

$$t_{\rm adv} \simeq \frac{R_{\rm diss}/\Gamma}{c} \sim \frac{R_{\rm diss}\theta}{c} \simeq 4.9\times10^5\,\epsilon^{-1}_{\rm rad,\text{-}1}\,\Gamma_1^2\theta_{\text{-}1}M_{\bullet,9}\ {\rm s}\ , \quad (28)$$

where we have made use of the fact that in the comoving frame the dissipation region is roughly spherical ($R_{\rm diss}/\Gamma \sim R_{\rm diss}\theta$). This is of the same order as the adiabatic cooling time due to the jet lateral expansion:

$$t_{\rm ad} \simeq \frac{1}{\beta_{\rm exp}}\frac{R_{\rm diss}\theta}{c} \simeq 4.9\times10^5\,\epsilon^{-1}_{\rm rad,\text{-}1}\,\Gamma_1\,M_{\bullet,9}\ {\rm s}\ , \quad (29)$$

where $\beta_{\rm exp} \simeq \Gamma\theta$ is the jet expansion velocity in the comoving frame.

As Table 1 shows, advection and adiabatic cooling are the fastest loss processes for 1-10 GeV ions. The table also includes the characteristic timescales for the following processes: radiative cooling via synchrotron and inverse Compton emission, inelastic collisions with other protons ($p\,p$ collisions, with cross section $\sigma_{pp} \simeq 3\times10^{-26}$ cm^2 at energies of a few GeV and inelasticity $k_{pp} \simeq 1/2$) and inelastic collisions with background photons ($p\,\gamma$ collisions) – resulting in photo-meson production and/or pair production via the Bethe-Heitler effect (for an estimate of the corresponding cross section, see Sikora et al. 1987). We also discuss the limiting case in which the proton thermal energy stored in the dissipation region (see $U'_{\rm th}$ in eq. (14)) is efficiently transferred to the emitting electrons, which then cool via synchrotron and inverse Compton at the observed bolometric luminosity $L_{\rm Bol} \sim 10^{48}$ erg s^{-1}; the corresponding timescale is a conservative lower limit which holds regardless of uncertainties in the efficacy of energy exchange between protons and electrons.

TABLE 1
Important Timescales for $1-10$ GeV Protons in the Jet Dissipation Region

PHYSICAL PROCESS	TIMESCALE (s)	SYMBOL
Shock Heating	$\mathbf{3.9\times10^{-5}}\,\gamma_p$	$t_{\rm heat}$
Fermi Acceleration	$2.1\times10^2\,\gamma_p^{1/3}$	$t^{\rm (Kolm)}_{\rm DSA}$
	6.9×10^4	$t^{\rm (str)}_{\rm DSA}$
Advection	$\mathbf{4.9\times10^5}$	$t_{\rm adv}$
Adiabatic Cooling	$\mathbf{4.9\times10^5}$	$t_{\rm ad}$
Synchrotron and IC Cooling	$8.8\times10^{15}\,\gamma_p^{-1}$	
$p-p$ Inelastic Collisions	1.1×10^{11}	
$p-\gamma$ Inelastic Collisions	$2.4\times10^{12}\,\gamma_p^{-1}$	
$p\rightarrow e$ Energy Transfer	1.7×10^6	
Diffusion	$1.1\times10^9\,\gamma_p^{-1/3}$	$t^{\rm (Kolm)}_{\rm diff,J}$
	$\mathbf{3.5\times10^6}$	$t^{\rm (str)}_{\rm diff,J}$

When evaluating the above timescales, we use $\Lambda_{\rm Edd}=1$, $\Gamma=10$, $\theta=0.1$, $\epsilon_{\rm rad}=0.1$, $\epsilon_{\rm th}=0.1$, $\epsilon_{\rm B}=0.1$ and $M_\bullet=10^9\,M_\odot$. We also assume that $L_{\rm IC}/L_{\rm S}=10$. In boldface, the timescales most relevant for our model.

4.1.2. *Cosmic Ray Diffusion out of the Jet*

In the tangled and inhomogeneous fields of the dissipation region, a fraction of the shock-heated protons may efficiently scatter with resonant magnetic fluctuations and diffuse out of the jet before being advected away with the jet flow. The cosmic ray diffusion timescale across the jet is given by

$$t_{\rm diff,J} \simeq \tau_{\rm J}\frac{R_{\rm diss}\theta}{c} \sim \tau_{\rm J}\,t_{\rm adv}\ , \quad (30)$$

where $\tau_{\rm J}$ is the cosmic ray optical depth.[5] The resonant magnetic fluctuations providing the cosmic ray scattering may be embedded in the jet flow with, e.g., a Kolmogorov spectrum, or generated in the dissipation region by the accelerated cosmic rays themselves via the so called "streaming instability" (e.g., Kulsrud & Pearce 1969; Wentzel 1974).

A concise review of both scattering processes and their ability to describe the cosmic ray distribution function in the Milky Way is given in §3 of Socrates et al. (2008). We assume that the scattering mechanisms which account for the interstellar cosmic ray optical depth in the Galaxy operate in the core of quasar jets as well and we employ the same scalings as in Socrates et al. (2008). This may seem as a huge extrapolation, but in the absence of direct observational constraints on the cores of radio-loud quasars, we utilize this assumption for lack of a better choice. We now estimate the value of $\tau_{\rm J}$ expected for the two scattering processes mentioned above.

The cosmic ray optical depth in the presence of resonant Kolmogorov turbulence at the Larmor scale of shock-heated protons ($r_{\rm L,J}=c/\omega_{\rm L}$) may be written

$$\tau_{\rm J}^{\rm (Kolm)} \simeq \left(\frac{\delta B_0}{B'}\right)^2\left(\frac{R_{\rm diss}\theta}{r_{\rm L,J}}\right)^{1/3}\left(\frac{R_{\rm diss}\theta}{\lambda_0}\right)^{2/3}\ , \quad (31)$$

where λ_0, δB_0 and B' are respectively the stirring scale of the magnetic turbulence, the fluctuation amplitude at the stirring scale and the jet magnetic field computed in eq. (16). Assuming $\delta B_0 \sim B'$ and $\lambda_0 \sim R_{\rm diss}\theta$, eq. (31) gives

$$\tau_{\rm J}^{\rm (Kolm)} \simeq 2.3\times10^3\,\gamma_p^{-1/3}\Lambda_{\rm Edd}^{1/6}\epsilon_{\rm B,\text{-}1}^{1/6}\epsilon_{\rm th,\text{-}1}^{1/6}\Gamma_1^{-1/3}M_{\bullet,9}^{1/6}\ , \quad (32)$$

and the corresponding diffusion time across the jet is

$$t_{\rm diff,J}^{\rm (Kolm)} \simeq 1.1\times10^9\gamma_p^{-1/3}\Lambda_{\rm Edd}^{1/6}\epsilon_{\rm B,\text{-}1}^{1/6}\epsilon_{\rm th,\text{-}1}^{1/6}\epsilon_{\rm rad,\text{-}1}^{-1}\Gamma_1^{5/3}\theta_{\text{-}1}M_{\bullet,9}^{7/6}\ {\rm s}\ . \quad (33)$$

In the case that the cosmic ray opacity is provided by magnetic fluctuations self-generated via the streaming instability, the optical depth may be written

$$\tau_{\rm J}^{\rm (str)} \simeq \frac{c}{v_{A,\rm J}} \simeq 7.1\,\epsilon_{\rm B,\text{-}1}^{-1/2}\epsilon_{\rm th,\text{-}1}^{-1/2}\ , \quad (34)$$

where $v_{A,\rm J} = B'/\sqrt{4\pi\rho'_{\rm diss}}$ is the Alfvén velocity in the dissipation region. The diffusion time across the jet is

[4] Diffusive shock acceleration occurs on a timescale $t_{\rm DSA} \sim t_{\rm adv}/\tau_{\rm J}$, where $\tau_{\rm J} \gtrsim 1$ is the cosmic ray optical depth in the jet dissipation region (see §4.1.2). Therefore, although not required in our model, there would be enough time to accelerate a fraction of the shock-heated protons to suprathermal energies before the flow is advected away from the shock.

[5] The hierarchy between the characteristic timescales for diffusion ($t_{\rm diff,J}$), advection ($t_{\rm adv}$) and diffusive shock acceleration ($t_{\rm DSA}$) is such that $t_{\rm diff,J} \sim \tau_{\rm J}\,t_{\rm adv} \sim \tau_{\rm J}^2\,t_{\rm DSA}$, and since $\tau_{\rm J} \gtrsim 1$ we have $t_{\rm diff,J} \gtrsim t_{\rm adv} \gtrsim t_{\rm DSA}$.

then

$$t_{\rm diff,J}^{\rm (str)} \simeq 3.5\times10^{6}\,\epsilon_{\rm B,-1}^{-1/2}\,\epsilon_{\rm th,-1}^{-1/2}\,\epsilon_{\rm rad,-1}^{-1}\,\Gamma_1^2\,\theta_{-1}\,M_{\bullet,9}\ {\rm s}\ . \quad (35)$$

By modeling the spectral energy distribution of powerful blazars, Celotti & Ghisellini (2008) infer that the $\sim 0.5-10$ GeV electrons responsible for the observed synchrotron and inverse Compton emission should be injected at internal shocks with a power-law energy spectrum $n(\gamma_e)\propto\gamma_e^{-\alpha}$ with $\alpha\simeq2.5$. If the mechanism responsible for the electron acceleration to suprathermal energies depends only on particle rigidity, as is the case for a Fermi process, then a comparable spectral index should describe the accelerated protons. For first-order Fermi acceleration at mildly-relativistic internal shocks this implies, for $\sim$ GeV protons, that $\tau_{\rm J}\Delta E/E\sim\tau_{\rm J}\Delta\Gamma/\Gamma\sim1$, since the inferred particle spectrum has nearly equal energy per logarithmic interval. Here, $\Delta E/E$ is the average fractional energy gain per acceleration cycle. It follows that $\tau_{\rm J}\sim\Gamma/\Delta\Gamma\gtrsim1$.

4.2. *Cosmic Ray Efficiency $\epsilon_{\rm CR}$ and Luminosity $L_{\rm CR}$*

Armed with a better knowledge of the physical processes relevant for 1-10 GeV protons produced in the jet core, we can now compute the fraction $\epsilon_{\rm CR}$ of jet power injected into the host galaxy as cosmic rays, and the corresponding interstellar cosmic ray luminosity $L_{\rm CR}$.

Since the fraction of shock-heated protons that can diffuse out of the jet before advection takes place is $\sim t_{\rm adv}/t_{\rm diff,J}\sim1/\tau_{\rm J}$, as shown in eq. (30), the cosmic ray efficiency in eq. (26) may be rewritten as $\epsilon_{\rm CR}\simeq\epsilon_{\rm th}/\tau_{\rm J}$. If the cosmic ray scattering is due to resonant magnetic fluctuations with a Kolmogorov spectrum (cf. eq. (32)), then

$$\epsilon_{\rm CR}^{\rm (Kolm)}\simeq\frac{\epsilon_{\rm th}}{\tau_{\rm J}^{\rm (Kolm)}}\simeq4.4\times10^{-5}\,\gamma_p^{1/3}\,\frac{\epsilon_{\rm th,-1}^{5/6}\,\Gamma_1^{1/3}}{\Lambda_{\rm Edd}^{1/6}\,\epsilon_{\rm B,-1}^{1/6}}\,M_{\bullet,9}^{-1/6}\ ,\quad(36)$$

where $\gamma_p\sim1-10$ may be thought of as the characteristic comoving Lorentz factor of shock-accelerated protons. Instead, if the cosmic ray opacity is provided by magnetic waves self-generated via the streaming instability (cf. eq. (34)), the resulting cosmic ray efficiency is

$$\epsilon_{\rm CR}^{\rm (str)}\simeq\frac{\epsilon_{\rm th}}{\tau_{\rm J}^{\rm (str)}}\simeq1.4\times10^{-2}\,\epsilon_{\rm B,-1}^{1/2}\,\epsilon_{\rm th,-1}^{3/2}\ . \quad (37)$$

As eqs. (36) and (37) suggest, the cosmic ray coupling efficiency $\epsilon_{\rm CR}=L_{\rm CR}/L_{\rm J}$ is either weakly dependent (for Kolmogorov turbulence) or not dependent at all (for the streaming instability) on the black hole mass $M_\bullet$. In other words, our cosmic ray-driven feedback scenario is manifestly *self-similar* or *scale-independent per unit of jet power* (or equivalently, per unit of black hole mass, for fixed radiative efficiency).

The cosmic ray luminosity $L_{\rm CR}=\epsilon_{\rm CR}L_{\rm J}$ injected into the host galaxy is then, for the two scattering mechanisms discussed above,

$$L_{\rm CR}^{\rm (Kolm)}\simeq5.7\times10^{42}\gamma_p^{1/3}\,\frac{\Lambda_{\rm Edd}^{5/6}\epsilon_{\rm th,-1}^{5/6}\,\Gamma_1^{1/3}}{\epsilon_{\rm B,-1}^{1/6}}\,M_{\bullet,9}^{5/6}\ {\rm erg\,s^{-1}}\ ,\quad(38)$$

$$L_{\rm CR}^{\rm (str)}\simeq1.8\times10^{45}\,\Lambda_{\rm Edd}\,\epsilon_{\rm B,-1}^{1/2}\,\epsilon_{\rm th,-1}^{3/2}\,M_{\bullet,9}\ {\rm erg\,s^{-1}}\ ,\quad(39)$$

where we have made use of $L_{\rm J}=\Lambda_{\rm Edd}L_{\rm Edd,\bullet}$.

Comparison of eq. (39) with the Eddington limit in cosmic rays defined in eq. (4) (or equivalently, of eq. (37) with eq. (7)) suggests that, if the resonant magnetic turbulence for cosmic ray scattering in the jet core is mainly generated by the streaming instability, there may be enough power in interstellar cosmic rays to launch a momentum-driven wind which could eject the galactic gas. Although the *momentum* requirement $L_{\rm CR}\gtrsim L_{\rm Edd,CR}$ is a necessary condition for efficient self-regulation of the black hole – galaxy system, it is not sufficient by itself; in fact, we must also require that the *time-integrated* cosmic ray *energy* injected into the host galaxy during this epoch of super-Eddington cosmic ray activity should be comparable to the binding energy of the gas. In §5 we address this important issue.

4.3. *Cosmic Ray Propagation within the Host Galaxy*

In the frame of the host galaxy, the shock-accelerated protons which diffuse out of the jet are relativistically beamed along the jet axis within an angle $\sim1/\Gamma$. Since the cosmic ray optical depth for scattering off resonant magnetic fluctuations in the galaxy's interstellar medium is very large ($\tau_g\sim10^3$, as we show below), the cosmic ray momentum distribution quickly isotropizes so that their overall effect on the galactic gas resembles a spherically-symmetric pressure force in the direction opposite to the galaxy's gravitational center. As the interstellar cosmic rays diffuse towards larger scales, their energy does not appreciably change from the value $\simeq\Gamma\gamma_p m_p c^2\sim1-100$ GeV of their birth, since the average fractional energy loss per scattering is only $\sim\tau_g^{-2}\ll1$ (e.g., Kulsrud & Pearce 1969; Wentzel 1974). Instead, the fractional momentum change per scattering is $\sim\tau_g^{-1}$, which allows cosmic rays to be extremely effective in powering a momentum-driven outflow of interstellar gas. In this section, we study the cosmic ray propagation within the host galaxy, providing an estimate for the cosmic ray optical depth τ_g and diffusion time $t_{{\rm diff},g}$.

The characteristic galactic scale radius R_g for an isothermal sphere is

$$R_g\simeq\frac{G}{2}\frac{M_\star}{\sigma_\star^2}\simeq18.2\left(\frac{M_\star/M_\bullet}{10^3}\right)\left(\frac{M_{\bullet,9}}{0.76\,\sigma_{\star,300}^4}\right)\sigma_{\star,300}^2\ {\rm kpc}\ ,\quad(40)$$

if the black hole – galaxy system follows the $M_\bullet-M_\star$ and $M_\bullet-\sigma_\star$ relations. We take R_g as the characteristic distance where cosmic rays produced in the jet core interact with the galactic gas.

If we pin the ratio between cosmic ray mean free path and Thomson mean free path at the galactic scale R_g to the value $\lambda_{\rm CR}/\lambda_{\rm T}\sim10^{-6}$ appropriate for the Milky Way, the cosmic ray optical depth up to R_g is

$$\tau_g\simeq\kappa_{es}\,\rho_g R_g\frac{\lambda_{\rm T}}{\lambda_{\rm CR}}\simeq1.5\times10^3 f_{g,-1}\left(\frac{\lambda_{\rm T}/\lambda_{\rm CR}}{10^6}\right)\ ,\quad(41)$$

where ρ_g is the gas mass density at R_g for an isothermal sphere. The cosmic ray diffusion time up to R_g is then

$$t_{{\rm diff},g}\simeq\tau_g\frac{R_g}{c}\simeq8.9\times10^7 f_{g,-1}\left(\frac{\lambda_{\rm T}/\lambda_{\rm CR}}{10^6}\right)\sigma_{\star,300}^2\ {\rm yr}\ .\quad(42)$$

We expect that, during epochs of vigorous jet-powered cosmic ray activity, the galactic structure will be appreciably affected by the cosmic ray pressure on timescales comparable to the cosmic ray diffusion time computed in eq. (42). It follows that, if the radio jet is powered by gas accretion onto the central black hole, then $t_{\rm diff,\it g}$ might be a reasonable upper limit for the duration of the radio-loud phase. In §5.2 we further comment on this and examine the implications of the scaling $t_{\rm diff,\it g} \propto \sigma_\star^2$.

As discussed by Socrates et al. (2008), the cosmic ray pressure force may be reduced by $\sqrt{t_{pp,g}/t_{\rm diff,\it g}}$ if the diffusion timescale $t_{\rm diff,\it g}$ is significantly longer than the time $t_{pp,g}$ required to deplete the cosmic ray energy via inelastic scattering with background protons resulting in pion production. For a cross section $\sigma_{pp} \simeq 3 \times 10^{-26}$ cm^2 at $\sim$ GeV energies and a scattering inelasticity $\kappa_{pp} \simeq 1/2$, we find

$$t_{pp,g} \simeq \frac{m_p}{\kappa_{pp}\sigma_{pp}c}\frac{1}{\rho_g} \simeq 1.8 \times 10^8\, f_{g,\text{-}1}^{-1}\, \sigma_{\star,300}^2 \text{ yr} , \quad (43)$$

which is marginally longer than $t_{\rm diff,\it g}$, so that we can ignore losses due to pion production.[6]

Finally, we show that cosmic ray optical depths comparable to the value in eq. (41) may be derived by making the extreme assumption that the magnetic energy density in the galaxy is roughly in equipartition with the gas random kinetic energy density, i.e. $B^2/8\pi \sim 3/2\, \rho_g \sigma_\star^2$ at the characteristic radius R_g. If the cosmic ray scattering is provided by background magnetic turbulence with a Kolmogorov spectrum, the optical depth up to R_g is

$$\tau_g^{\rm (Kolm)} \simeq \left(\frac{\delta B_0}{B}\right)^2 \left(\frac{R_g}{r_{{\rm L},g}}\right)^{1/3} \left(\frac{R_g}{\lambda_0}\right)^{2/3} \simeq 4.4{\times}10^3 f_{g,\text{-}1}^{1/6}\, \sigma_{\star,300}^{2/3} , \quad (44)$$

where $\lambda_0 \sim R_g$ is the stirring scale and $\delta B_0 \sim B$ the fluctuation amplitude at the stirring scale for the interstellar magnetic turbulence, and $r_{{\rm L},g}$ is the Larmor radius in the galactic magnetic field B for a $\sim$ 10 GeV proton. Instead, if the cosmic ray opacity results from Alfvén waves self-generated via the streaming instability, the cosmic ray optical depth will be

$$\tau_g^{\rm (str)} \simeq \frac{c}{v_{A,g}} \simeq 5.8 \times 10^2 \sigma_{\star,300}^{-1} , \quad (45)$$

where $v_{A,g} = B/\sqrt{4\pi\rho_g}$ is the Alfvén velocity in the galaxy. Note that both scattering mechanisms yield an optical depth comparable to the result in eq. (41).

[6] We point out that cosmic ray losses due to pp collisions are negligible all the way down to pc scales. If the Thomson optical depth in the BLR is $\tau_{\rm T} \sim 0.1$, the extreme assumption that the cosmic ray diffusion time in the BLR be comparable or longer than the proton-proton collision time implies a CR optical depth at pc scales $\tau_{\rm CR} \gtrsim m_p\kappa_{es}/(\kappa_{pp}\sigma_{pp})\, \tau_{\rm T}^{-1} \simeq 450$. In the BLR, the gravitational force is still primarily provided by the black hole, and the correspoding Eddington luminosity in CRs will be $L_{\rm Edd,CR} = L_{\rm Edd,\bullet}\, \tau_{\rm T}/\tau_{\rm CR} \simeq 2.0 \times 10^{43} M_{\bullet,9}$ erg s^{-1}. This is smaller than the CR luminosity in eq. (39). Therefore, in this case, the CR flux would disrupt the entire BLR. Since the BLR *exists*, we can comfortably reject the extreme assumption that CRs are significantly destroyed at pc scales.

5. COSMIC RAY FEEDBACK RESULTING FROM EPISODIC RADIO-LOUD ACTIVITY: ORIGIN OF THE $M_\bullet - M_\star$ RELATION

As discussed in §4.2, during epochs of powerful jet activity the cosmic ray luminosity $L_{\rm CR}$ injected into the interstellar medium of the host galaxy may exceed the galaxy's Eddington limit in cosmic rays $L_{\rm Edd,CR}$ (compare eqs. (39) and (4)). Due to the outward cosmic ray pressure force, hydrostatic balance is lost and a cosmic ray momentum-driven wind develops, removing the fuel for further star formation and black hole accretion. However, the bulge and black hole growth is completely choked off only if the radio-loud phase lasts long enough such that the total energy output in cosmic rays, which would eventually couple to the interstellar medium, is comparable to the binding energy of the galactic gaseous component. In other words, the *momentum* requirement $L_{\rm CR} \gtrsim L_{\rm Edd,CR}$ is a necessary prerequisite, but efficient self-regulation of the black hole – galaxy system is achieved only when the *energy* requirement discussed above is fulfilled as well.

A generic feedback mechanism is capable of self-regulating the combined black hole – galaxy growth only if the total energy injected into the interstellar medium $\Delta E_{\rm inj}$ is comparable to the gravitational energy of the gas:

$$\Delta E_{\rm inj} \sim E_g \simeq f_g M_\star \sigma_\star^2 , \quad (46)$$

where we have adopted an isothermal sphere. In what follows, we assume that the gas fraction f_g is independent of black hole mass and stellar velocity dispersion. We now contrast our jet-powered cosmic ray feedback model with self-regulation mechanisms that act during the optically-luminous "quasar phase."

5.1. *Black Hole Self-Regulation during the Luminous Quasar Phase*

The Soltan (1982) argument, along with the work of Yu & Tremaine (2002), indicates that supermassive black holes at the center of galaxies build up their mass primarily by an act of radiatively-efficient accretion, in a relatively short-lived high-luminosity "quasar phase". During this epoch, the accretion luminosity approaches the black hole Thomson Eddington limit $L_{\rm Edd,\bullet}$.[7]

If energy release during the optically-bright quasar phase is responsible for black hole self-regulation, the total energy $\Delta E_{\rm QP}$ absorbed by the galactic gas during the quasar lifetime $\Delta t_{\rm QP}$ should satisfy $\Delta E_{\rm QP} \sim E_g$, as prescribed by eq. (46). If only a fraction $\epsilon_{\rm QP}$ of the accretion luminosity $L_{\rm acc}$ can couple to the interstellar medium of

[7] Caution must be taken when applying the Soltan (1982) argument to black hole demographics, encapsulated by the Magorrian et al. (1998) relation. The luminosity and mass function of supermassive black holes peak close to a black hole mass $M_\bullet \sim 10^8\, M_\odot$ (see dotted line in Fig. 1). Relatively small black holes, like the one at the center of the Milky Way, are both too rare and faint, such that a Soltan-like argument cannot be employed in order to determine whether or not their mass was built up by radiatively-efficient Eddington-limited accretion. However, it seems likely that relatively small black holes with masses as low as $M_\bullet \simeq 10^6 - 10^7\, M_\odot$ can radiate close to the Eddington limit, which therefore suggests that the build-up of black hole mass in relatively small systems takes place during a bright short-lived "quasar phase" as well.

the host galaxy, we require

$$\Delta E_{\rm QP} \simeq \epsilon_{\rm QP}\, \Delta t_{\rm QP}\, L_{\rm acc} \sim E_g \ . \quad (47)$$

Assuming $L_{\rm acc} \sim L_{\rm Edd,\bullet}$, this can be rewritten as

$$\frac{M_\bullet}{M_\star} \sim \left(\frac{\sigma_{\rm T}}{4\pi\, G\, m_p\, c}\right) \frac{f_g\, \sigma_\star^2}{\epsilon_{\rm QP}\, \Delta t_{\rm QP}} \ . \quad (48)$$

The quasar lifetime $\Delta t_{\rm QS}$, set by the Soltan (1982) argument, is roughly comparable to the Salpeter time of the black hole:

$$t_{\rm Salp} = \frac{\Delta E_\bullet}{L_{\rm Edd,\bullet}} \simeq \left(\frac{\sigma_{\rm T}\, c}{4\pi\, G\, m_p}\right) \epsilon_{\rm rad} \simeq 4.5 \times 10^7 \epsilon_{\rm rad,\text{-}1}\ {\rm yr}\ , \quad (49)$$

where $\Delta E_\bullet \simeq \epsilon_{\rm rad} M_\bullet c^2$ is the total radiative energy output of the accretion process. Since $\epsilon_{\rm rad}$ depends only on black hole spin, the quasar lifetime $\Delta t_{\rm QS} \sim t_{\rm Salp}$ should be independent of black hole mass. It follows from eq. (48) that, in order for a system to lie on the $M_\bullet - M_\star$ relation, the coupling efficiency $\epsilon_{\rm QP}$ in the luminous quasar phase must depend on the stellar velocity dispersion $\sigma_\star$ such that

$$\epsilon_{\rm QP} \propto \sigma_\star^2 \ . \quad (50)$$

This implies that, if the quasar phase were to self-regulate the black hole and galaxy evolution, the feedback mechanism at work would *not* be a *universal*, *scale-free* or *self-similar* process. We stress again that this follows from $\Delta t_{\rm QP} \sim t_{\rm Salp}$, or equivalently from the fact that the total quasar energy output $\Delta t_{\rm QP} L_{\rm acc}$ is observationally pinned to be close to the maximum amount of energy released in the formation of the black hole, given by $\Delta E_\bullet$.[8]

There is no apparent reason why the coupling efficiency $\epsilon_{\rm QP}$ may not depend on the stellar velocity dispersion $\sigma_\star$ or the black hole mass $M_\bullet$. However, it should depend upon $\sigma_\star$ and $M_\bullet$ in such a way that makes the ratio $M_\bullet/M_\star$ a constant across nearly four decades in black hole mass. This is, of course, not impossible, but it would have to involve a cosmic conspiracy with respect to the gas dynamics of black hole self-regulation. We now show how our jet-powered cosmic ray feedback scenario provides a satisfactory solution for this apparent contradiction.

5.2. *Black Hole Self-Regulation During the Explosive Radio-Loud Phase*

For a generic feedback mechanism operating during epochs of AGN radio-loud activity, the total energy $\Delta E_{\rm RP}$ injected into the interstellar medium is

$$\Delta E_{\rm RP} \simeq \epsilon_{\rm RP} \Delta E_{\rm J} \simeq \epsilon_{\rm RP} \Delta t_{\rm RP} L_{\rm J}\ , \quad (51)$$

where $\Delta E_{\rm J} \simeq \Delta t_{\rm RP} L_{\rm J}$ is the time-integrated kinetic output of the radio jet, $\Delta t_{\rm RP}$ is the duration of the radio-loud phase and $\epsilon_{\rm RP}$ is the efficiency for coupling the jet power $L_{\rm J}$ to the galactic gas. The energy balance $\Delta E_{\rm RP} \sim E_g$ in eq. (46) requires that

$$\Delta E_{\rm RP} \simeq \epsilon_{\rm RP}\, \Delta t_{\rm RP}\, \Lambda_{\rm Edd} L_{\rm Edd,\bullet} \sim E_g\ , \quad (52)$$

where $\Lambda_{\rm Edd} = L_{\rm J}/L_{\rm Edd,\bullet}$. This can be rewritten as

$$\frac{M_\bullet}{M_\star} \sim \left(\frac{\sigma_{\rm T}}{4\pi\, G\, m_p\, c}\right) \frac{f_g\, \sigma_\star^2}{\epsilon_{\rm RP}\, \Delta t_{\rm RP}\, \Lambda_{\rm Edd}}\ , \quad (53)$$

from which it is apparent that the product $\epsilon_{\rm RP}\, \Delta t_{\rm RP}\, \Lambda_{\rm Edd}$ for radio-loud phases replaces $\epsilon_{\rm QP}\, \Delta t_{\rm QP}$ in quasar epochs (cf. eq. (48)). For a black hole – galaxy system that lies on the $M_\bullet - M_\star$ relation, this implies

$$\epsilon_{\rm RP} \Delta t_{\rm RP}\, \Lambda_{\rm Edd} \propto \sigma_\star^2\ . \quad (54)$$

As opposite to the quasar feedback scenario discussed in §5.1, eq. (54) does not immediately constrain the coupling efficiency $\epsilon_{\rm RP}$ to depend upon the stellar velocity dispersion. The possibility that the feedback efficiency $\epsilon_{\rm RP}$ is independent of $M_\bullet$ and $\sigma_\star$ would then require that $\Delta t_{\rm RP}\, \Lambda_{\rm Edd}$ increases with the stellar velocity dispersion, or that the time-integrated jet kinetic output per unit of black hole mass should scale as $\propto \sigma_\star^2$.

If black hole self-regulation in the radio-loud phase is mediated by cosmic rays produced in the jet core, we have shown that the coupling efficiency $\epsilon_{\rm CR}$ is indeed independent (for cosmic rays scattering with Alfvén waves self-generated via the streaming instability, see eq. (37)) or weakly dependent (for resonant diffusion by magnetic turbulence with a Kolmogorov spectrum, see eq. (36)) on the black hole mass.[9] In our cosmic ray-driven feedback model we should then expect that $\Delta t_{\rm RP}\, \Lambda_{\rm Edd} \propto \sigma_\star^2$ or, if $\Lambda_{\rm Edd}$ is a constant with black hole mass, that the duration of the radio-loud phase should scale as $\Delta t_{\rm RP} \propto \sigma_\star^2$.

As discussed in §4.3, a reasonable upper limit for $\Delta t_{\rm RP}$ may be given by the cosmic ray diffusion timescale $t_{{\rm diff},g}$ within the host galaxy. In fact, if the jet is powered by black hole accretion, AGN radio activity would be terminated when the interstellar cosmic rays have diffused up to the scales where most of the gas resides, and started to push it outward, thus preventing further accretion. Interestingly, if f_g and $\lambda_{\rm CR}/\lambda_{\rm T}$ do not depend on black hole mass, eq. (42) shows that $t_{{\rm diff},g} \propto \sigma_\star^2$, the same scaling that $\Delta t_{\rm RP}$ should have in order to satisfy eq. (54) with a scale-independent coupling efficiency.

Note that for $\Delta t_{\rm RP} \sim t_{{\rm diff},g}$, eq. (42) suggests that the duration of the radio-loud phase for the most massive galaxies may be comparable to the Salpeter time in eq. (49), whereas in relatively small galaxies it should last less than the quasar phase. In other words, for low-mass galaxies the black hole self-regulation would be confined to an episodic radio-loud epoch much shorter than the time required to accrue the black hole mass.

[8] The scaling in eq. (50) seems to be in contradiction with the simulations by Di Matteo et al. (2005), which reproduce the $M_\bullet - \sigma_\star$ relation assuming that a constant fraction of black hole accretion energy is deposited in the galactic gas (Springel et al. 2005). However, we remark that $\epsilon_{\rm QP}$ in eq. (50) is the fraction of accretion energy *available to unbind the galactic gas*, and not just the fraction of energy *deposited in the gas*. If the simulations by Di Matteo et al. (2005) are taken at face value, then there is a hidden parameter that limits the fraction of deposited energy which is available to unbind the galactic gas. The hidden parameter should scale as $\propto \sigma_\star^2$.

[9] We have defined $\epsilon_{\rm CR}$ as the fraction of jet power injected into the host galaxy as cosmic rays. However, due to the large cosmic ray optical depth within the host galaxy (see §4.3), the whole interstellar cosmic ray power will eventually be transferred to the galactic gas, so that $\epsilon_{\rm CR}$ is also a good proxy for the feedback coupling efficiency.

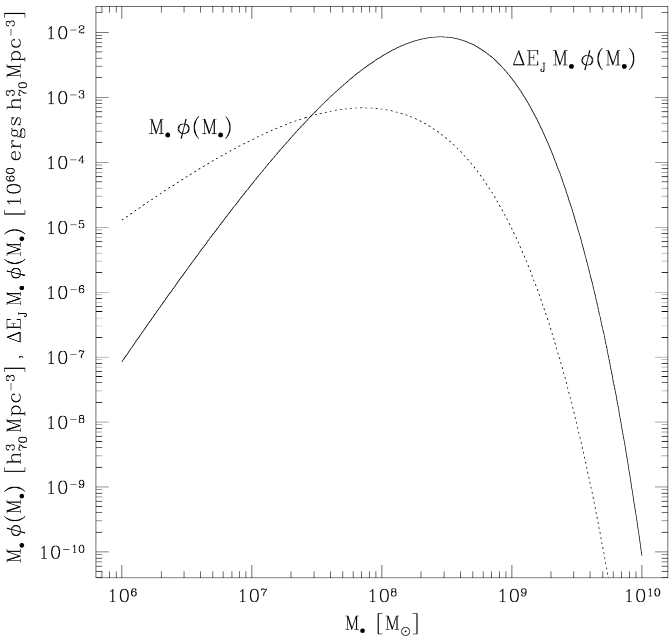

FIG. 1.— Black-hole mass function ($M_\bullet\, \phi(M_\bullet)$; dotted line) and kinetic energy output of radio jets integrated over cosmic time ($\Delta E_{\rm J}\, M_\bullet\, \phi(M_\bullet)$; solid line), versus black hole mass $M_\bullet$. The function $\phi(M_\bullet)$, namely the number density of black holes with mass in the interval $[M_\bullet, M_\bullet + {\rm d}M_\bullet]$, is derived from the stellar velocity dispersion function of early-type galaxies by Sheth et al. (2003) via the $M_\bullet - \sigma_\star$ relation. The jet time-integrated kinetic output $\Delta E_{\rm J}$ is supposed to satisfy the energy balance $\epsilon_{\rm RP} \Delta E_{\rm J} \sim f_g M_\star \sigma_\star^2$ required for black hole self-regulation during radio-loud phases, as in eq. (52). We have assumed a coupling efficiency $\epsilon_{\rm RP} = 10^{-3}$, as found for our cosmic ray-driven feedback model (somewhat an intermediate value between eq. (36) and eq. (37)) and a gas fraction $f_g = 0.1$ for the host galaxy; the black hole – galaxy system lies on the $M_\bullet - M_\star$ and $M_\bullet - \sigma_\star$ relations. We set h_{70} to be the Hubble constant H_0 in units of 70 km s^{-1} Mpc^{-1}.

6. DISCUSSION AND COMPARISON WITH OBSERVATIONS

We propose a self-regulation mechanism for supermassive black holes and their host bulges that operates in actively-accreting systems with powerful radio-loud activity. In the core of relativistic radio jets, internal shocks resulting from the dissipative interaction of shells of matter, intermittently ejected from the central engine, convert a fraction of the ordered kinetic energy of the jet flow into thermal form. If a sufficient number of the protons (or "cosmic rays") heated and randomized at internal shocks escape the radio core into the interstellar medium of the host galaxy, they may profoundly affect the evolution of the galaxy and its black hole. To quantify their effect on the hydrostatic balance of the galactic gas we define an Eddington limit in cosmic rays for the host galaxy. For a phenomenologically-motivated jet model, we show that during powerful radio-loud phases the power in 1-100 GeV interstellar cosmic rays is large enough to break the cosmic ray Eddington limit for the host galaxy, so that *momentum* balance of the galactic gaseous component is lost and a cosmic ray-driven wind develops, that removes the galactic gas. If this super-Eddington cosmic ray activity lasts long enough, the time-integrated cosmic ray *energy* input into the interstellar medium may exceed the binding energy of the gas, and the whole galaxy's gaseous phase will become unbound. In doing so, any fuel for further black hole growth and star formation will be removed, thus affecting the combined evolution of the black hole and its stellar bulge, as implied by the $M_\bullet - M_\star$ and $M_\bullet - \sigma_\star$ relations.

Morganti et al. (2005) report the detection of *fast* (~ 1000 km s^{-1}) *large-scale* ($\sim 1 - 10$ kpc) massive outflows of *neutral* hydrogen in several powerful radio galaxies. They claim that the outflows are spatially associated with radio knots extended along the jet, but this may just result from the fact that H I absorption can be traced only in the presence of a bright background radio continuum. Fast large-scale outflows of low-ionization species with nearly *spherical* morphology have been detected in the powerful radio source MRC 1138-262 by Nesvadba et al. (2006). The fact that the outflow is "cold" (atomic or weakly ionized) and "fast" (super-virial) implies that it is the result of catastrophic momentum, rather than energy, exchange with the interstellar medium. In other words, it is possible that some sort of Eddington limit is being broken. Among the momentum-driven feedback mechanisms, Nesvadba et al. (2006) conclude that neither radiation-powered quasar winds (e.g., Fabian 1999; King 2003; Murray et al. 2005) nor direct coupling between the jet and the interstellar medium in the "dentist drill" model (Begelman & Cioffi 1989) seem sufficient to explain the large-scale gas kinematics observed in MRC 1138-262. Direct coupling between the jet and the galactic gas should result in the outflow being confined within a narrow cone around the jet axis, contrary to observations. Regarding the effect of radiation pressure from the quasar photon output on interstellar dust grains (e.g., Murray et al. 2005), it is hard to understand how this could remove a significant fraction of the galactic gas, since the UV photons will be degraded to IR frequencies after a few scatterings, within a parsec from the black hole. Also, the optical depth for the resulting far IR photons will not exceed unity beyond a few tens of parsecs.

The jet-powered cosmic ray feedback scenario presented here does not suffer any of these deficiencies. The cosmic ray luminosity injected into the host galaxy is smaller than the black hole photon power by roughly two orders of magnitude, but cosmic rays exchange momentum with the galactic gas $\sim \tau_g/\tau_{\rm UV} \sim 10^3$ more efficiently. Here, $\tau_{\rm UV} \sim 1$ is the optical depth for UV photons on dust grains. Moreover, since the cosmic ray optical depth in the galaxy's interstellar medium is very large ($\tau_g \sim 10^3$), any memory of the cosmic ray momentum distribution at injection is quickly lost and their effect on the galactic gas should resemble a spherically-symmetric outward pressure force, so that the resulting gas outflow is nearly spherical. Also, the fraction of cosmic ray energy lost at each interaction with the interstellar gas is minimal ($\sim \tau_g^{-2}$), meaning that they can propagate up to the large scales where most of the gas resides without suffering significant losses.

An interesting result of our study is that, per unit of black hole energy release, the cosmic ray feedback efficiency $\epsilon_{\rm CR}$, which gives the fraction of jet power available to unbind the galactic gas, is roughly a constant with black hole mass. Instead, for self-regulation models relying on the black hole photon output during radiatively-efficient quasar phases (e.g., Silk & Rees 1998; Fabian

1999; Ciotti & Ostriker 2001; King 2003; Di Matteo et al. 2005; Murray et al. 2005; Hopkins et al. 2006), the feedback efficiency is constrained to scale as $\sim E_g/\Delta E_\bullet \propto \sigma_\star^2$ in order for the system to lie on the $M_\bullet - M_\star$ relation. Here, $E_g \simeq f_g M_\star \sigma_\star^2$ is the binding energy of the galactic gas and $\Delta E_\bullet$ is the time-integrated energy output resulting from black hole accretion. It would be a surprising coincidence if an intrinsically *scale-dependent* self-regulation mechanism were to result in the *scale-free* $M_\bullet - M_\star$ relation, which holds for nearly four decades in mass.

The explosive radio-loud phase does not suffer from such severe constraints. Rather, since the coupling efficiency $\epsilon_{\rm CR}$ for our cosmic ray-driven feedback scenario is scale-independent, the energy balance $\epsilon_{\rm CR}\Delta E_{\rm J} \sim E_g$ required for black hole self-regulation implies that the time-integrated jet kinetic output should scale as $\Delta E_{\rm J} \propto M_\star \sigma_\star^2$ for a constant gas fraction f_g. By inferring the black hole mass function $M_\bullet\,\phi(M_\bullet)$ (dotted line in Fig. 1) from the velocity dispersion function of early-type galaxies by Sheth et al. (2003) via the $M_\bullet - \sigma_\star$ relation, we can predict the dependence on black hole mass of the kinetic energy output of radio jets integrated over cosmic time ($\Delta E_{\rm J} M_\bullet\,\phi(M_\bullet)$; solid line in Fig. 1), for systems lying on the $M_\bullet - M_\star$ and $M_\bullet - \sigma_\star$ relations. We find that its slope at the low-mass end is ~ 2.5 and it peaks at $M_\bullet \simeq 3\times10^8\,M_\odot$. With more reliable measurements of jet kinetic power and black hole mass, this could provide a stringent observational test for our proposed self-regulation scenario. For example, finding a slope close to 2.5 at the low-mass end would imply that the ratio between the time-integrated jet kinetic output and the binding energy of the galactic gas should be independent of black hole mass, thus strongly implicating that black hole self-regulation occurs in the radio-loud phase, irrespective of the actual coupling mechanism itself.

For a generic feedback mechanism acting during radio-loud epochs, the scaling $\Delta E_{\rm J} \propto M_\star \sigma_\star^2$ can be recast as $\Delta t_{\rm RP}\Lambda_{\rm Edd} \propto \sigma_\star^2$, for a black hole – galaxy system which follows the $M_\bullet - M_\star$ relation. Here, $\Lambda_{\rm Edd}$ is the ratio of the jet kinetic power to the black hole Thomson Eddington limit and $\Delta t_{\rm RP}$ is the duration of the radio-loud phase. This implies that, for systems where a significant amount of mass is being built up, radio-loud signatures will most likely be observed in galaxies with large black holes and bulges, in line with the basic phenomenology of radio-loud AGN and quasars (e.g., Laor 2000; Best et al. 2005). The $M_\bullet - M_\star$ relation perhaps is then connected to the fact that radio-loud actively-accreting objects are relatively absent in relatively small systems (Greene et al. 2006): black hole self-regulation takes place during the radio-loud phase and, due to the abundance of total energy available, this phase is short-lived in relatively small systems because the amount of energy required to unbind the galaxy is relatively low.

We thank Annalisa Celotti, Peter Goldreich, Jenny Greene, Jerry Ostriker and Anatoly Spitkovsky for helpful discussions. AS acknowledges support from a Lyman Spitzer Jr. Fellowship given by the Department of Astrophysical Sciences at Princeton University and a Friends of the Institute Fellowship at the Institute for Advanced Study in Princeton.

REFERENCES

Antonucci, R. 1993, ARA&A, 31, 473
Begelman, M. C. & Cioffi, D. F. 1989, ApJ, 345, L21
Best, P. N. et al., 2005, MNRAS, 362, 25
Bîrzan, L., Rafferty, D. A., McNamara, B. R., Wise, M. W., & Nulsen, P. E. J. 2004, ApJ, 607, 800
Celotti, A. & Ghisellini, G. 2008, MNRAS, 203
Ciotti, L. & Ostriker, J. P. 2001, ApJ, 551, 131
Di Matteo, T., Springel, V., & Hernquist, L. 2005, Nature, 433, 604
Fabian, A. C. 1999, MNRAS, 308, L39
Ferrarese, L. & Merritt, D. 2000, ApJ, 539, L9
Gebhardt, K. et al., 2000, ApJ, 539, L13
Greene, J. E., Ho, L. C., & Ulvestad, J. S. 2006, ApJ, 636, 56
Häring, N. & Rix, H.-W. 2004, ApJ, 604, L89
Ho, L. C. 2002, ApJ, 564, 120
Ho, L. C. & Peng, C. Y. 2001, ApJ, 555, 650
Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2006, ApJS, 163, 1
Jorstad, S. G. et al., 2005, AJ, 130, 1418
King, A. 2003, ApJ, 596, L27
Kobayashi, S., Piran, T., & Sari, R. 1997, ApJ, 490, 92
Kulsrud, R. & Pearce, W. P. 1969, ApJ, 156, 445
Laor, A. 2000, ApJ, 543, L111
Lind, K. R. & Blandford, R. D. 1985, ApJ, 295, 358
Magorrian, J. et al., 1998, AJ, 115, 2285
Marconi, A. & Hunt, L. K. 2003, ApJ, 589, L21
McLure, R. J. & Dunlop, J. S. 2002, MNRAS, 331, 795
Morganti, R., Tadhunter, C. N., & Oosterloo, T. A. 2005, A&A, 444, L9
Murray, N., Quataert, E., & Thompson, T. A. 2005, ApJ, 618, 569
Nesvadba, N. P. H., Lehnert, M. D., Eisenhauer, F., Gilbert, A., Tecza, M., & Abuter, R. 2006, ApJ, 650, 693
Rawlings, S. & Saunders, R. 1991, Nature, 349, 138
Sckopke, N., Paschmann, G., Brinca, A. L., Carlson, C. W., & Luehr, H. 1990, J. Geophys. Res., 95, 6337
Shakura, N. I. & Sunyaev, R. A. 1973, A&A, 24, 337
Sheth, R. K. et al., 2003, ApJ, 594, 225
Shurkin, K., Dunn, R. J. H., Gentile, G., Taylor, G. B., & Allen, S. W. 2008, MNRAS, 383, 923
Sikora, M., Kirk, J. G., Begelman, M. C., & Schneider, P. 1987, ApJ, 320, L81
Sikora, M., Madejski, G., Moderski, R., & Poutanen, J. 1997, ApJ, 484, 108
Silk, J. & Rees, M. J. 1998, A&A, 331, L1
Socrates, A., Davis, S. W., & Ramirez-Ruiz, E. 2008, ApJ, 687, 202
Soltan, A. 1982, MNRAS, 200, 115
Springel, V., Di Matteo, T., & Hernquist, L. 2005, MNRAS, 361, 776
Tanaka, M., Goodrich, C. C., Winske, D., & Papadopoulos, K. 1983, J. Geophys. Res., 88, 3046
Tremaine, S. et al., 2002, ApJ, 574, 740
Urry, C. M. & Padovani, P. 1995, PASP, 107, 803
Wentzel, D. G. 1974, ARA&A, 12, 71
Yu, Q. & Tremaine, S. 2002, MNRAS, 335, 965